Mapping of Terrorist Activities in Kenya using Sentiment Analysis

Lucas .A. Ngoge
Strathmore University
Faculty of Information Technology
Nairobi, Kenya
Lucas.ngoge@strathmore.edu

Joseph Onderi Orero
Strathmore University
Faculty of Information Technology
Nairobi, Kenya
jorero@strathmore.edu

***Abstract*— Terrorism has become a subject of concern to many people in Kenya today. Corruption, porous border and lack of government in the neighboring Somali, have made Kenya a potential target for terrorists'.**

The advancement in technology has brought a new era in criminal activities where Online Social Networks such as Twitter, Facebook has driven the increase use of the internet by criminal organizations and their supporters for a wide range of purposes including recruitment, financing, propaganda, incitement and gathering and dissemination of information for criminal activities such as threats, incitement to imminent violence, harassing speech, libelous speech etc.

Although the Kenya government improved its ability to fight terrorism but the changing pattern of terrorist activities, human errors and delayed crime analyses have given criminals more time to destroy evidence and escape arrest.

The main objective is to test and validate a technique that can be used to establish crime patterns associated with terrorist activities using sentiment information deduced from twitter data. The data collected will then be used as features for training and development of the algorithm which will then be used to carry out real time mapping of terrorist activities. The algorithm's performance will be then measured for accuracy.

Keywords— Machine Learning,Sentiment Analysis, Corpus, Mapping, Classification.

I. INTRODUCTION

Kenya is considered to be a beacon of stability and peace in the horn of Africa, greater Eastern Africa region and in Africa, Kenya plays an important role. It serves as an economic and business hub for both national and international investors (Leah & Abdalla, 2009). However, terrorism has threatened its peace and stability. It has undermined the freedom of association and movement of citizens and created a sense of fear and intimidation which has hampered the spiritual, economic and social development of individuals (Leah & Abdalla, 2009).

Terrorism has been propagated using Technology, corruption, porous border, lack of government in the neighboring Somali amongst others, have immensely made Kenya a potential target for terrorists' (Samuel, 2013). The Terrorists are using technology to recruit, finance, spread propaganda, gather and disseminate information for terrorist activities (Jytte, 2015).

The advancement in technology has brought a new era in terrorism where Online Social Networks such as Twitter, Facebook has driven the increase use of the internet by terrorist organizations and their supporters for a wide range of purposes. Twitter has recently emerged as terrorists' favorite Internet service, even more popular than self-designed websites or Facebook, to disseminate propaganda and enable internal communication. For example, without Twitter, the explosive growth of ISIS over the last few years into the most-feared terrorist group in the world would not have been possible (Weimann, 2014). In addition, the militant group Al-Shabaab, during its September 2013 attack on Westgate Mall in Nairobi, Kenya, gave a live commentary on its actions on Twitter. While terrorists have developed many ways to use the Internet in furtherance of illicit purposes, their use of the Internet also provides opportunities for the gathering of intelligence and other activities to prevent and counter acts of terrorism, as well as for the gathering of evidence for the prosecution of such acts (UN, 2012).

The means by which the internet is often utilized to promote and support acts of terrorism and criminal activities are propaganda (recruitment, radicalization, and incitement to terrorism), financing, training and planning acts of terrorism (UN, 2012).
First, Propaganda; it takes the forms of multimedia communications providing ideologies or practical instructions or explanations or justifications for promotion of terrorist activities.
Second, financing; terrorists use the Internet to raise and collect funds and resources (UN, 2012).
Third, training; internet provides a platform for detailed instructions, often easily accessible multimedia format and multiple languages, on topics such as how to join clandestine organizations, how to construct explosives, firearms or other weapons or hazardous materials and how to plan and execute terrorist attacks (UN, 2012).
Finally, planning; it is an act of terrorism typically involves remote communication among several parties. Internet technology facilitates the preparation of acts of terrorism, through communications within and between organizations promoting violent extremism, as well as across borders (UN, 2012).

The increased use of internet for criminal purposes provides a corresponding increase in the availability of electronic data which may be compiled and analyzed for counter-terrorism purposes. Law enforcement, intelligence and other authorities are developing increasingly sophisticated tools to proactively prevent, detect and deter terrorist activity involving use of the Internet (UN, 2012). Although the Kenya government improved its ability to fight terrorism by increasing its capabilities to identify, arrest and detain suspects through an Anti-Terrorism Police Unit (ATPU), but the changing pattern of terrorist activities, human errors and delayed crime analyses have given criminals more time to destroy evidence and escape arrest (Hsuchun et al, 2004).

Therefore, sentiment analysis and mapping can be used to help with identification, detection and tracking terrorists' activities in an actionable and timely manner.

II. LITERATURE REVIEW

A. Crime pattern analysis and mapping

Crime analysis is a law enforcement function that involves the systematic analysis of identifying and analyzing both pattern and trends in crime and disorder (Bolla, Raja & Ashok, 2014). Criminal pattern analysis is very crucial in combating crime.

Computer systems have to be engaged in order to gather and interpret intelligence so as to control the criminal environment as well as influence effective decision making (Kester, Quist-aphetsi & Mieee, 2013). Figure 1 shows stages in criminal analysis. Analysis of a crime pattern typically focuses on who, what, when, where, and how factors that are common across a significant number of incidents are identified. Identification of these commonalities is often the key to finding solutions to criminal activities (IACA, 2008).

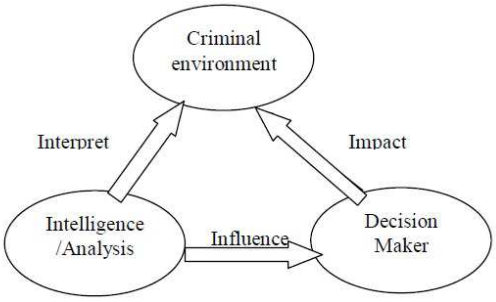

Figure 1: Crime analysis process

The crime analysis process begins with collecting and managing data thereafter the actual analysis of a crime pattern begins with its initial identification of criminal activities and their patterns. To accomplish crime analysis, the analysts must ensure that he has real-time access to incident reports and make information scanning a daily process (Kester, Quist-aphetsi & Mieee, 2013).

The analyst can determine the criminal patterns based on one of three factors: first, the modus operandi commonalities, found through a careful review of incident reports and their narratives; second, Exceptional Volume, found through some brand of threshold analysis, either deliberate or unconscious; and lastly, Geographic Proximity, found through crime mapping.

i. Crime Mapping

The crime mapping method of pattern identification involves creating pin maps or thematic maps of various crimes, and seeking geographic hot spots or clusters. This method is valid only for crime patterns that exhibit geographic clustering (Kester, Quist-aphetsi & Mieee, 2013).

Many crime patterns involve crimes that are not geographically close, so, like threshold analysis, the crime mapping method should not be the only means by which an analyst seeks patterns. However, like threshold analysis, mapping crime may help catch patterns lost in the detail of daily report review, and it may be beneficial for understaffed agencies that cannot effectively use daily report review.

B. *Automated techniques for detection of online activities*

Automated techniques aim at identifying patterns from structured and unstructured data for example collected twitter data. Some of the common methods for crime data mining being used are; entity extraction, clustering, deviation detection, classification amongst others.

i. Entity extraction

This method seeks to identify a particular pattern from text. It is used to identify people, vehicles, and addresses from police narrative reports. The investigators can use this method to analyze the behaviors of serial offenders.

However, this method provides very basic information for crime analysis because its performance relies on the availability of large amount of input data (Hsuchun et al, 2004). An example of an application that uses this method is the advanced terrorist detection system (ATDS). This application tracks down terrorist-generated sites, by analyzing the websites visited by the users. Figure 2 and figure 3 show how ATDS works.

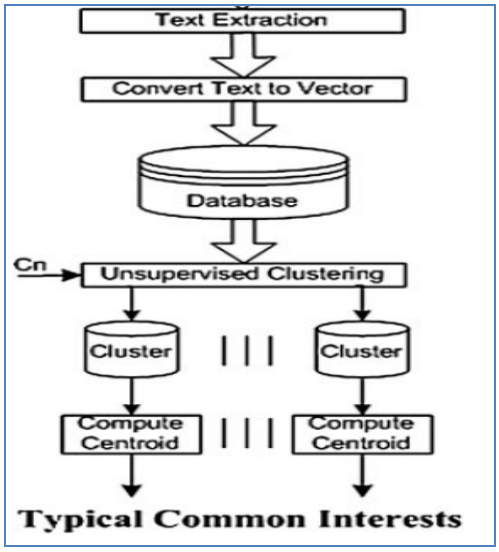

Figure 2: Learning mode

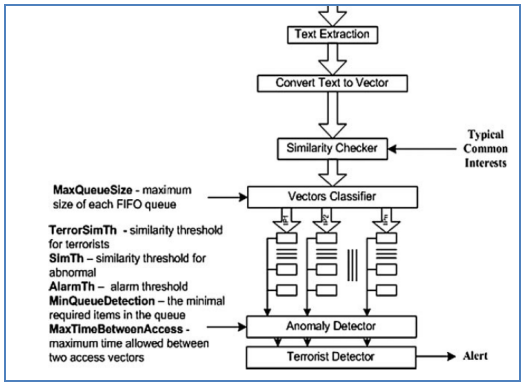

Figure 3: Detection mode

Once the learning mode (in figure 2) is completed, the system can switch to the detection mode. In the detection mode (in figure 3), the system detects users who are accessing typical content by intercepting HTML web pages. Each new incoming intercepted web

page is converted into a vector of weighted terms and similarity is computed between the current vector and the known group profiles. The detection algorithm makes a decision whether the user has recently visited the site or a series of pages by considering the history of a user's visit to the site.

If in the learning mode, a profile of typical terrorist content has been built, then the system will also check the similarity of abnormal users to terrorist content. In the detection mode the pages that users access is captured, filtered, and transferred to a vector representation then the clustering module accesses the collected vectors and stores in the database and performs unsupervised clustering (e.g., using k-means algorithm) resulting in n clusters. The number of clusters Cn is one of the system parameters specified by the users (Yuval elovici et al, 2007).

ii. Deviation detection

This method uses specific measures to study data that differs markedly from the rest of the data. An example of an application that uses this method is the outlier detection system. Outlier detection system identifies and finds patterns in data that do not conform to expected behavior.

Outliers or anomalies can be detected without knowing the data set's distribution or needing any labelled training samples (Svetlana C, 2005). In the context of credit card fraud detection, a fraudulent transaction can be seen as an outlier which behaves differently comparing to legitimate transactions hence it can easily be spotted (Svetlana C, 2005).

Outliers sometimes called Anomalies are patterns in data that do not conform to a well-defined notion of normal behavior (Varun, Arindam, & Vipin, 2009). Figure 4 illustrates anomalies in a simple 2-dimensional data set. The data has two normal regions, N1 and N2, since most observations lie in these two regions. Points that are sufficiently far away from the regions, e.g., points o1 and o2, and points in region O3, are anomalies.

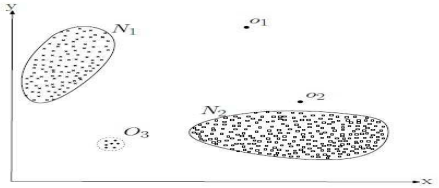

Figure 4: A simple example of anomalies in a 2-dimensional data set

Anomalies might be induced in the data for a variety of reasons, such as malicious activity, e.g., credit card fraud, cyber-intrusion, terrorist activity or breakdown of a system, but all of the reasons have a common characteristic that are interesting to the analyst (Varun, Arindam, & Vipin, 2009). Figure 5 shows the above mentioned key components associated with any anomaly detection technique.

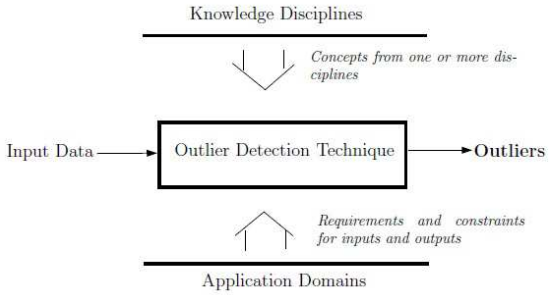

Figure 5: components of outliers' detection technique

This technique has extensive use in a wide variety of applications such as military surveillance for terrorist activities. The law enforcement officers can apply this method to detect fraud, crime analyses and network intrusion.

However, some of these activities that this method guard against may appear to be normal making it difficult for this method to detect (Hsuchun et al, 2004).

iii. Classification

Classification tries to find common features amongst different crime entities and organizes them into predefined classes. It is commonly used to predict crime patterns because it lessens time to identify crime entities. The investigators can use this method to reveal the identity of cyber criminals who use internet to spread radical information to the public.

However, this technique requires a predefined classification scheme and a large amount training data and testing data because if there will be any missing data from the training set then accuracy of prediction would be limited thereby making it unsuitable for crime prediction (Hsuchun et al ,2004).

C. Sentiment analysis

Sentiment analysis is the study that analyses people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, and their attributes. Sentiment analysis requires training document of textual content or a data corpus, which serves as a preparation document for classification.

The basic machine learning techniques available for text classifications are naive bayes (NB), support vector machines (SVM), maximum entropy amongst others (Pang & Lee, 2008). The law enforcement officers can use these machine learning techniques to predict terrorist activities by analyzing factors such as time, location, address, physical characteristics etc. from a body of a text.

i. Sentiment analysis processes

The sentiment analysis process involves data collection, text preparation, sentiment detection, sentiment classification and presentation of output (Pang & Lee, 2008).

a. Data Collection

The data collection involves the collection of tweets according to some crime related topics for example use of keywords to identify text/tweets. The keywords such as," Gun", "crime", "kill", "Allah", "kaffir", "attack", "blast", "bomb", "claim", "defend", "force", "kill",

"protect", "resist", "strike", etc. is used to track tweets from the twitter (Pang & Lee, 2008).

Tweet Dataset is a brief details of dataset used in experiments. Dataset needs to be cleaned and tokenized and should be in a CSV format with the headers such as id, date, tweet and location. An example of tweet Dataset is shown in the table 1 below.

ID	Date	Tweet	Loc	Lat	Lon
12112	Tue Mar 08 16:23:23 +0000 2016	US Bombs Al-Shabaab Terror Camps Graduation ;killing over 150 Terrorists .Https://t.coKn7AMJFrNF	Embu	-0.425	37.531

Table 1: Tweet Dataset Format

b. Text preparation

Text preparation involves removal of any unwanted text from tweets collected from twitter and convert them into lower case before classification is performed i.e. removal of alphanumeric characters i.e. hash, dashes (Pang & Lee, 2008). Table 2 shows a sample of unwanted text.

Table 2: Sample of unwanted content and action

S/No.	Unwanted Conten	Wanted Content	Action
1.	#word	Word	Replaced
2.	@username	AT_USER	Converted
3.	https://	URL	Converted
4.	Additional White Space ' '	" "	Removed
5.	Retweet	RT	Removed
6.	Uppercase	lowercase	Converted

c. Sentiment detection

Sentiment detection requires that each sentence is examined for subjectivity using unigrams and bigrams thereafter a geographical analysis is to be conducted that is, collecting and scrutinizing every sample in a set of samples from which samples are drawn (Pang & Lee, 2008).

d. Sentiment classification

It involves running a sentiment classifier on each extracted sentence to determine if it is positive, negative, or neutral. The basic classifiers available for text classification are naives Bayes, support vector machine, maximum entropy amongst others (Pang & Lee, 2008).

D. Machine learning techniques for text classification

Machine learning is the study of the algorithms that are capable in fully automated situations to predict something out of input. They are, Naive Bayes, support vector machines (SVM), maximum entropy etc.

Many prototypes and models for sentiment classification treat classifiers and feature extractors as two distinct components (Pang & Lee, 2008).

i Naive Bayes

(Pang &Lee, 2008) describes naive Bayes classifier as a supervised machine learning algorithm with a simple probabilistic classifier based on bayes theorem with strong independence assumptions.

However, the classifier assumes the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of the other feature.

ii support vector machines (SVM)

SVM is a supervised learning model which analyzes the data and identifies the pattern for classification. In text classification data is linearly divisible and produces very accurate classifications however, it is computationally expensive (Pravesh & Mohd ,2014).

iii Maximum entropy

The principle in maximum entropy is that when nothing is known, the distribution should be as uniform as possible, that is, have maximal entropy. Labeled training data is used to derive a set of constraints for the model that is characterized by the class-specific expectations for the distribution.

Using maximum entropy model, prediction of outcome is based on everything that is known and assumes nothing about unknown (Govindarajan & Romina, 2013).

E. Lexicon-Based Approaches

Lexicon based method uses sentiment dictionary with opinion words and match them with the data to determine polarity (Vishal & Sonawane, 2016).

They assign sentiment scores to the opinion words describing how positive, negative and objective the words contained in the dictionary.

i. Dictionary-based

It is based on the usage of terms (seeds) that are usually collected and annotated manually. This set grows by searching the synonyms and antonyms of a dictionary (Vishal & Sonawane, 2016).

ii. Corpus-Based

The corpus-based approach provides the dictionaries related to a specific domain. These dictionaries are generated from a set of seed opinion terms that grows through the search of related words by means of either statistical or semantic techniques (Vishal & Sonawane,2016).

iii. Bag of words model

Bag of words approach is a method of document classification where the frequency of each word is used as feature for training and developing a classifier. A text is represented as a bag (multi set) of words disregarding grammar and even words order but keeping multiplicity (Pang & Lee, 2008).

Bag of words technique uses unigram features to carry out sentiment analysis and classification on twitter data. N-gram is technique of finding n-grams words from a given document (Pang & Lee, 2008). The commonly used model for this techniques includes unigram (n=1), bigrams (n=2) and trigrams (n=3). Unigram consists of all individual words present in the text; bigrams defines a pair adjacent words. Each pair words form a single bigram (Pang & Lee, 2008).

F. Conceptual model

The summary of the design and architecture of the model is given by the figure 6 below. The figure clearly shows the relationships between individual elements that make the model. Bag of words technique was used to develop the model.

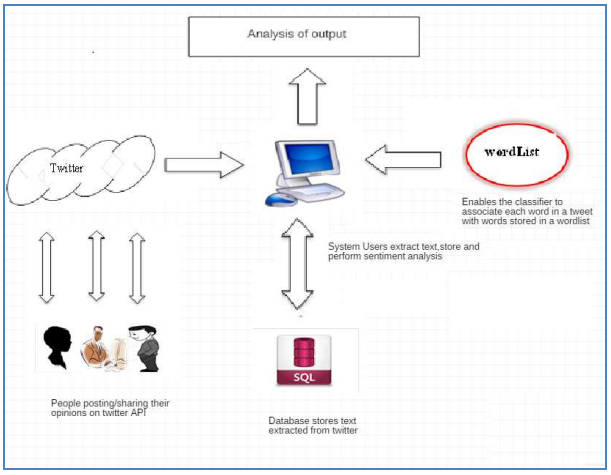

Figure 6: Conceptual model

III. SYSTEM DESIGN AND ARCHITECTURE

a. Systems Design

Systems design is the dissection of a system into its component pieces for purposes of studying how those component pieces interact and work according to (Gemino & Parker, 2009).

The sequence diagram in figure 7 shows the functional and non-functional requirements of the system. The components in the sequence diagram are; twitter, which represent the twitter server, the user, and the end-user who is focused to use the information in the database.

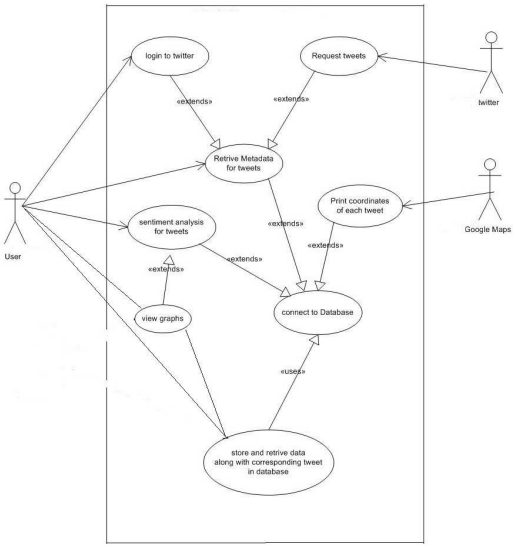

Figure 7: System components

b. System architecture

The figure 8 shows the components that make up the system. The components present the tasks that a user can carry out using the system. These tasks are; data collection, data cleansing, classification, analysis and mapping of sentiment onto a map

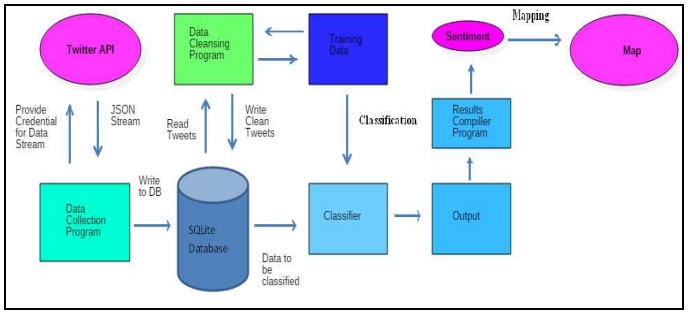

Figure 8: Structure of the system

c. Classification of sentiments

This is a process whereby data collected from Twitter is classified into three classes' namely negative, positive and neutral class. A classifier was created to read tweets from the database which involved building of a database or CSV files of both positive and negative wordlists.

A tweet is represented as a bag of words which is broken down into individual words. Each word is matched to the words stored in the positive and negative wordlist. When there is a match the counter is incremented or decremented by a fixed number depending on the weight or value assigned to each word in the wordlist.

When this process is completed the classifier classifies the tweet as positive, negative or neutral.

d. Mapping of sentiments

This is a process of linking classified tweets and its geo-coordinates on to the map as shown fig 9. The map is produced with markers distributed across various points on it so that the sentiments can be easily visualized by the user.

The three classes are indicated on the map using different colors i.e. red for negative, green for positive and yellow for neutral sentiment as shown in fig 10,11 and 12.

```
Positive Tweet,Kericho,-0.273,35.383,2
Neutral Tweet,Nandi,0.055,35.193,0
Negative Tweet,Nandi,0.055,35.193,-1
Negative Tweet,Mandera,3.36667,40.7,-1
Neutral Tweet,Vihiga,0.072,34.712,0
Positive Tweet,Isiolo,0.98333,38.53333,1
Neutral Tweet,Machakos,-1.282,37.408,0
Neutral Tweet,Busia,0.35,34.17,0
Neutral Tweet,Kakamega,0.334,34.797,0
Positive Tweet,Wajir,1.75,40.01667,1
Positive Tweet,Marsabit,2.96667,37.6,2
Positive Tweet,Isiolo,0.98333,38.53333,1
Neutral Tweet,Taita Taveta,-3.4,38.37,0
Neutral Tweet,Samburu,1.33333,37.11667,0
Positive Tweet,West Pokot,1.75,35.25,1
Neutral Tweet,Elegeyo-Marakwet,0.99,35.55,0
Positive Tweet,Homa Bay,-0.666,34.481,3
Neutral Tweet,Siaya,0.105,34.302,0
Neutral Tweet,Garissa,-0.172,40.041,0
Positive Tweet,Nairobi,-1.28333,36.83333,2
Neutral Tweet,Isiolo,0.98333,38.53333,0
```

Figure 9: A sample of classified data with locations

IV. TESTING AND VALIDATION

i. Testing

Ease of use and clarity of the system is tested to ensure that the system meets the requirements of the users.

ii. Test results

The test results from the classifier when evaluated using a dataset containing 346 tweets collected from twitter streaming API is displayed in table 3.

Table 3: Overall sentiment

S/No.	Instances	Total
1.	Total Number of instances	346
2.	Total instances of positive class	20
3.	Total instances of negative class	275
4.	Total instances of neutral class	51
5.	Percentage of instances classified positive	5.80%
6.	Percentage of instances classified negative	79.50%
7.	Percentage of instances classified neutral	14.70%

iii. Visualization

The visualization of the sentiment collected from twitter API is displayed and presented by the map in figure 10,11 and 12.

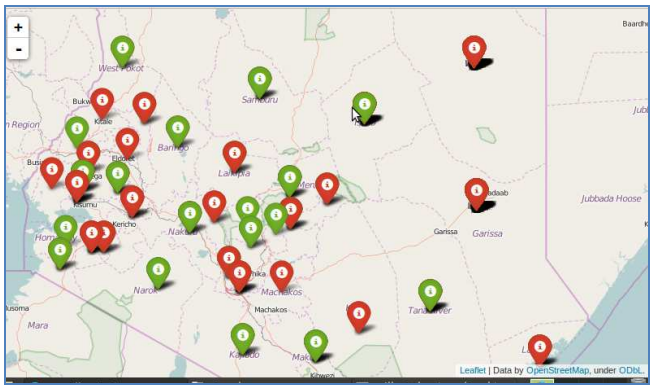

Figure 10: Sentiment distribution1

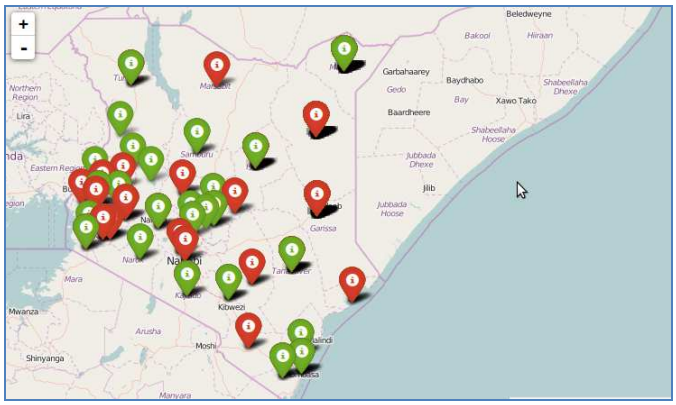

Figure 11: Sentiment distribution2

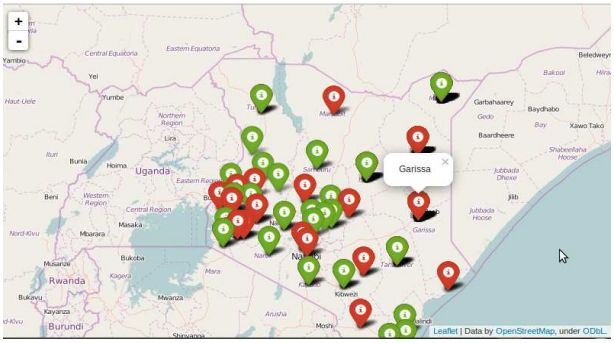

Figure 12: Location of the tweeter 1

The results show that there was high percentage of negative tweets related to terrorism compared to positive tweets. The locations where the sentiments originated were indicated on the map using markers i.e. red for negative, green for positive and yellow for neutral sentiment as shown in figure 11.

iv. Evaluation of the Model

The performance of sentiment classification can be measured by using the following equations; Accuracy = (TP+TN)/(TP+TN+FP+FN), Precision = TP/(TP+FP) and Recall = TP/(TP+FN). In which TP, FN, FP and TN refer respectively to the number of true positive instances, the number of false negative instances, the number of false positive instances and the number of true negative instances, as defined in the table 4 (Doaa,2016).

Table 4 Confusion Matrix

	Predicted Positives	Predicted Negatives
Actual Positive	TP	FN
Actual Negative	FP	TN

The true positive (TP) rate or recall is the rate at which positive text are predicted to be positive (R), whereas the true negative rate is the rate at which negative text are predicted to be negative.

The accuracy represents the rate at which the model predicts results correctly (A) (Doaa, 2016). The precision also called the positive predictive rate, calculates how close the measured values are to each other (P) (Doaa, 2016).

Table 5: Overall system performance

S/No.	Metric	Accuracy Score	
		Sentiment Analyzer	Naives Bayes
1.	Classification Accuracy	73%	60%
2.	Precision	60	50%

Table 5 shows a summary of the overall system performance; the accuracy level of the model was 73%, compared to naives bayes (60%) and the precision was (60%) against (50 %) of the naives bayes.

The system performance was high compared to naives bayes thus making it more efficient to be used in classifying real–time twitter data. The distribution of sentiments on the map as indicated by markers represented the patterns and trends of terrorist activities in Kenya

V. CONCLUSIONS

The research was guided by the three objectives. First was, to review the automated techniques for detection of online activities; to the extent of this objective, indeed a lot of researches have been done but they do not adequately address mapping of criminal activities using sentiments expressed in social media.

The second objective was to develop and implement a model that will help tracking and mapping of terrorist activities in Kenya; indeed, the system developed was able to map online activities at real time.

The third objective was to test and validate the model; the system performance was tested and its accuracy compared to naives bayes the results showed that the system performance was above 70% thus making it more efficient to be used in classifying real–time twitter data.

REFERENCES

[1] Doaa (2016), Enhancement Bag-of-Words Model for Solving the Challenges of Sentiment Analysis, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7/

[2] Govindarajan & Romina (2013), a survey of classification methods and applications for sentiment analysis, the international journal of engineering and science (IJES).

[3] Hsinchun et al (2004), A General Framework and Some Examples,

[4] Jytte (2015), tweeting the jihad, social media networks of western foreign fighters in Syria and iraq.

[5] Jeong-Yong & Aziz (2014), Internet of Things for Smart Crime Detection, Contemporary Engineering Sciences

[6] Kester, Quist-aphetsi & Mieee (2013), visualization and analysis of geographical crime patterns using formal concept analysis, international journal of remote sensing & geoscience (IJRSG), volume 2.

[7] Leah K & Abdallah B (2009), Social Policy, Development and Governance in Kenya.

Magutu P, Ondimu G & Ipu C (2011), Effects of Cybercrime on State Security: Types, Impact and Mitigations with the Fiber Optic Deployment in Kenya, Journal of Information Assurance & Cybersecurity

[8] National Crime Research Centre (NCRC) (2012), center summary of a study on organized criminal gangs in Kenya.

[9] Pang & Lee (2008), opinions mining and sentiment analysis. Foundations and trends in information extraction, vol 2

[10] Pravesh & Mohd (2014), methodological study of opinion mining and sentiment analysis techniques, international journal on soft computing (IJSC) vol. 5.

[11] Samuel (2013), neglecting history and geopolitics in approaches to counterterrorism, african journal of criminology and justice studies (AJCJS) vol.7.

[12] Svetlana C (2005), Outlier Detection in Clustering

[13] United Nation (UN) (2012), the use of the internet for terrorist purposes.

[14] Varun, Arindam, & Vipin (2009), anomaly detection, ACM computing surveys.

[15] Vishal & Sonawane (2016), sentiment analysis of twitter data: a survey of techniques, international journal of computer applications, volume 139

[16] Weimann (2014), New Terrorism and New Media. Washington, DC: Commons Lab of the Woodrow Wilson International Center for Scholars.

[17] Yuval elovici et al (2007), detection of access to terror-related web sites using an advanced terror detection system.